

Progressive Taxation and Redistribution

Pablo Beramendi[†]
Matthew Dimick[‡]
Daniel Stegmueller[§]

This version: October 2020

Abstract

In this paper we revisit the relationship between progressivity and redistribution. We develop a model that analyzes the relationship between the progressivity of income taxes, the design of fiscal systems, and the level of redistribution under different levels of income biased representation. Building on a novel measurement approach, that captures directly the direct (i.e. in the absence of behavioral responses) effect of taxes and benefits on incomes, we show three things: (1) Contrary to a widely shared view, the progressivity of taxes matters at least as much, if not more, than that of benefits; (2) There is a strong negative association between the relative importance of proportional taxes in the tax function and the progressivity of the income tax. This relationship determines the level of redistribution that can be achieved under different fiscal designs; and (3) As the political influence of the poor and the level of redistribution increase, the relative weight of proportional taxes in the fiscal system also increases.

Beramendi and Stegmueller's research was supported by the National Research Foundation of Korea (NRF-2017S1A3A2066657).

[†]Duke University, pb45@duke.edu

[‡]SUNY Buffalo Law School, mdimick@buffalo.edu

[§]Duke University, daniel.stegmueller@duke.edu

1. Introduction

With inequalities of wealth and income reaching levels unseen since the first Gilded Age (Alvaredo et al. 2018), the debate about which political tools are best suited to bring down inequality has been rekindled among politicians, pundits, and academics alike. Much of the debate about wealth taxation centers around the effectiveness of new instruments given the nearly unlimited ability of capital and wealth to seek refuge in friendlier jurisdictions and on the need to pursue new forms of international coordination among authorities with diverse, and often conflicting, incentives (Zucman 2015). At the same time, advanced societies have witnessed major changes in their distributions of labor and capital income. As Atkinson, Piketty, and Saez (2011) report, capital and labor income are both increasingly overlapping at the top end of the distribution. The slow development of coordinated forms of international capital taxation, the increasing levels of illicit financial flows and tax avoidance, and the partial merging of the distributions of labor and capital income all point to the need to revisit the scope of progressive income taxation as an income-equalizing tool in the short and medium run (Atkinson and Stiglitz 2015).

Given the potential for inefficiencies in redistributive interventions (Benabou 2002), what characteristics of the fiscal contract define the feasible levels of redistribution? What happens to the allocation of the burden across different income groups as the overall levels of redistribution change? And what are the implications of asymmetric levels of political influence by income groups?

To address these questions, we present a political economy model in which progressivity across rich democracies is a function of income biased representation. Within this setting, we analyze the relationship between the progressivity of the income tax and the overall reduction in the pre- and post-tax inequality. The model provides two major insights. First, as inequality increases, *both* tax progressivity and redistribution will decrease. The main intuition for this result is the increasing weight of the rich in the political process. Rich voters prefer regressive taxation, and as their influence increases, tax policy becomes more reflective of their preferences. One implication of the influence of the rich is that, although both tax progressivity and redistribution decrease with greater influence of the rich, progressivity and redistribution nevertheless move in the same direction. In other words, there is a positive association between progressivity and redistribution. Second, increases in the levels of redistribution, themselves reflective of a stronger political influence of the poor, lead to a sacrifice in the progressivity of the overall fiscal design. We uncover two reasons for this: (1) the poor prefer proportional taxes, relative to a more progressive tax structure, that maximizes revenue and increases transfers; the poor gain more from “soaking” the rich and the middle class than they lose in a proportional (i.e., more regressive) tax regime. And (2) middle-class income earners, compared to the poor, prefer a more progressive but less redistributive tax regime that shifts the tax burden from them to the very rich. It follows from both of these results that any system of representation that favors the preferences of the poor will produce a less progressive, but more redistributive, tax system than one that caters primarily to the middle class.

We provide compelling evidence on the two main implications of the model on the basis of a relatively underexploited technique among applied political economists. Our analyses build on a new measure of tax and benefit progressivity using a policy simulation approach. This approach

predicts the amount of tax a person would pay (or benefits received) based on all of the complex *statutory* legal rules specifying who should pay how much in taxes (or receive in transfers). This measures the “pure” effect of tax policy by excluding the ex-post effect of taxation on market—that is *pre-tax*—income.⁵ We uncover two main findings.

First, contrary to a widely held view, tax progressivity, as compared with social spending and benefit progressivity, is a substantively and statistically significant determinant of redistribution across rich democracies.⁶ We show empirically that more progressive tax structures are indeed associated with more redistribution, not less. On average, a standard deviation increase in tax progressivity increases redistribution by about 0.4 standard deviations. We also show that benefit progressivity has no clearly significant relationship to redistribution.

Second, we show that as the level of redistribution increases the fiscal system actually becomes more proportional and less progressive. In line with this reasoning, we also explore a corollary concerning political institutions: to the extent that political representation systems in advanced industrial societies facilitate the political incorporation of lower income groups, it is in these settings where the combination of high redistribution and lower overall progressivity of the fiscal system should become more apparent.

Our efforts relate to several previous lines of work. There is a large literature that tries to determine whether progressive tax structures will emerge under general voting equilibrium conditions (see Cukierman and Meltzer 1988; Snyder and Kramer 1988; Marhuenda and Ortuño-Ortín 1995; Roemer 1999; De Donder and Hindriks 2003, 2004; Carbonell-Nicolau and Klor 2003; Carbonell-Nicolau and Ok 2007; De Freitas 2009; Roemer 2011). Yet, this literature neither attempts to explain variation in progressivity across countries nor develops any systematic predictions about the relationship between progressivity and redistribution. We develop an explicit model of inequality, redistribution preferences, and biased political representation that addresses precisely these questions, while at the same time relaxing the non-regressivity assumptions in prior analyses of the relationship between political institutions and redistribution (e.g., Iversen and Soskice (2006)).

In addition, our argument and findings help resolve a fundamental ambiguity in previous contributions. In his seminal analyses, Kakwani (1977a, b) established that the overall redistributive incidence is a function of three factors: the size (effort) of fiscal spending, the progressivity of its design (of both taxes and benefits), and the reordering that both sets of policies cause through the behavioral responses of market actors in the distribution of market income in the first place. Leaving aside the latter term for now, it follows from this simple setup that, holding the level of effort constant, more progressive tax and benefit structures are more redistributive (Lambert

⁵In doing so, we adopt a rather draconian approach to the well know incidence problem (Atkinson and Stiglitz 2015). Most policies have second-order, behavioral effects, tax policy not least of all. Taxation does not change the distribution of income only by taking from some and giving to others. It also influences decisions such as how much to work, whether to work, how hard to bargain, when to pay taxes, and so forth. By doing so, taxation also influences the distribution of income before taxes are collected and redistributed. Measures of progressivity based on household income data conflate these two aspects (Beramendi 2001).

⁶Kenworthy (2008) makes both claims when he writes that among a set of developed countries, “[N]one . . . achieves much inequality reduction via taxes. Instead, to the extent inequality is reduced, it is mainly transfers that do the work.” Note that Kenworthy uses (descriptive) country comparisons, while our evidence is based on within-country changes.

2001). In other words, progressivity is a necessary condition for redistribution and there should be, empirically, a positive relationship between the two.⁷

Two streams of work in political science allegedly challenge this view. One simply argues that taxes have very limited (if at all) redistributive incidence. Rather, redistribution really takes place on the benefit side, not the tax side (e.g., Kenworthy 2008).⁸ The second line of work pushes this idea even further: the claim that countries that redistribute *more* have *less* progressive tax structures is quickly gaining widespread acceptance (see, e.g., Martin 2015; Kato 2003; Ganghof 2006; Beramendi and Rueda 2007; Martin, Mehrotra, and Prasad 2009; Prasad and Deng 2009; Martin 2015). It is important to note that most of this literature measures the incidence and distribution of taxes in terms of “effective taxes rates,” that is, the amount of taxes households actually pay as a share of household income. The goal to measure what households actually pay in taxes sounds like good, social-science sense. But as we will argue in just a moment, these measures ignore what is termed the behavioral incidence of taxes, which is what individuals or households *would have* earned (pre-tax) under a different set of *statutory* tax rates.

Our analysis helps clarify the tension between the theoretical importance of fiscal progressivity and the empirical study of the role of taxation in redistribution. The core of this tension derives from two sources. The first is a fundamental misunderstanding of how the incidence of various fiscal policies works. Taxes have a “mechanical” effect on income, by reducing the amount of household income between pre-tax and post-tax stages. But through “behavioral” channels, taxes also affect the level of pre-tax income, even before taxes are deducted.⁹ Measures of tax progressivity that ignore these effects (such as effective tax rates) will not actually measure the progressivity of the tax system. For example, an effective tax rate measure may make a highly-progressive statutory tax schedule look less so if higher top-tax rates have a sizeable effect on pre-market income inequality, because fewer people will actually be paying taxes at that top rate. Thus, without considering jointly in the same model and without making explicit assumptions about behavioral responses to taxation, it is simply not feasible to make claims about the distributional incidence of taxes versus spending policies (Atkinson and Stiglitz 2015).

The second is an ambiguity about the level of analysis on the basis of which different authors define progressivity. Some contributions focus on the relative balance *between* tax tools;¹⁰ others refer to progressivity *within* specific tax instruments (income, consumption, etc.).¹¹ Our model

⁷For a vivid illustration of the effect of taxes on distributive outcomes, see Newman and O'Brian (2011).

⁸This seems indeed an implicit prior in the comparative literature as a whole, where the focus on taxation has been relatively more limited in recent decades. An exception to this recent trend is the pioneering work of Alt (1983), Steinmo (1989), and Steinmo (1996).

⁹This can be because individuals, anticipating a lower return from their labor, substitute leisure for income and reduce their labor supply. In addition, taxes can affect how people, especially high-income earners, bargain over compensation with employers. Changes in tax rates may also encourage “tax avoidance” (Piketty, Saez, and Stantcheva 2014).

¹⁰It seems widely accepted that countries that redistribute more often use taxes that, all else equal, are less progressive. The main example here is the value-added tax (VAT), a type of consumption tax, which is used widely in (more redistributive) Europe, but not in the United States (see, e.g., Beramendi and Rueda 2007).

¹¹For example, exemptions on certain goods (e.g., food) can make consumption taxes more progressive. Moreover, the adoption of a consumption tax need not make the aggregate tax structure less progressive if, for example, the income tax schedule is made more progressive to compensate for the regressivity of the consumption tax.

incorporates both aspects in a common framework and studies how they respond to changes in overall fiscal effort.

The paper proceeds as follows. Section 2 develops an analytical framework to disentangle the relationship between progressivity and redistribution. The model results provide the basis for our empirical analyses (3). Subsequently, we detail how we measure tax progressivity and discuss our empirical strategy in section 4. In our results section (5) we demonstrate the relevance of tax progressivity (and the irrelevance of benefit progressivity) in determining the levels of redistribution achieved by countries and we subject our results to a variety of specification tests. We finish this section by providing an empirical illustration of further political implications from our model. Section 7 concludes the paper.

2. Model

In this section we develop a model to understand the relationship between progressivity and redistribution. After presenting the model set-up, we pay attention to two steps in the process: preference formation among three groups of voters (the rich, the poor, and middle income earners) and preference aggregation into policy.

2.1. Setup: Actors, tax function and redistribution

Consider a measure one continuum of individuals characterized by their marginal productivity w . There are three kinds of income earners, the rich, middle class, and poor, $i \in I = \{R, M, P\}$ with $w_R > w_M > w_P = 0$. Each group has density p_i with $\sum_{i \in I} p_i = 1$. We assume that $p_R, p_P < \frac{1}{2}$ so that the median individual has w_M .

Each individual pays a group-specific linear tax rate on income, y . Type- M individuals pay the rate t_M while type- R individuals pay t_R . (Since $y_P = 0$, $t_P = 0$.) Each individual also receives a lump sum transfer b . With the government budget constraint (see below), this makes the tax policy problem bidimensional: $\mathbf{t} = (t_1, t_2) \in T$, with $T : [0, 1] \times [0, 1]$. Formally, following standard definitions, the tax schedule will be (weakly) *progressive* whenever $T(y_i)/y_i$ is nondecreasing in income. Conversely, the tax schedule will be *regressive* whenever $T(y_i)/y_i$ is decreasing in income. Note that both progressive and regressive tax schedules are feasible within the set of possible tax schedules. In our case, a (weakly) progressive tax schedule simplifies to the condition $t_2 - t_1 \geq 0$ (or $t_2 - t_1 < 0$ for a regressive tax schedule).

We define redistribution in the standard way, as the condition where the distribution of disposable (post-tax, post-transfer) income Lorenz dominates the distribution of market (pre-tax, pre-transfer) income.¹² We can write this condition generally as:

$$\frac{\sum_{j=1}^k p_j c_j}{\bar{c}} - \frac{\sum_{j=1}^k p_j y_j}{\bar{y}} \geq 0 \quad \text{for each } k \in \{1, 2, \dots, n\} \quad \text{and } y_j > y_{j-1} \quad (1)$$

¹²Our empirical measure of redistribution is the difference in Gini coefficients of pre- and post-fisc income distributions. Since the Gini coefficient is simply measured by the area between the Lorenz curve and the line of equality, this definition of redistribution translates seamlessly to our empirical measure.

where y_i represents market income (with \bar{y} being average market income) and c_i is disposable income (with \bar{c} being average disposable income). These variables are further defined below. We require integer indexing to make clear that incomes are ordered in an increasing direction.

Because much of our paper is concerned about the relationship between tax progression and redistribution, it is important to clarify that one can have a tax schedule that is more progressive, but less redistributive, than another. Although progressivity in either the tax or transfer schedule (or both) is necessary for redistribution to occur, it is not sufficient. For instance, in the case where each tax rate taxes all income fully, $t_1 = t_2 = 1$, and transfers are lump sum, the tax schedule is proportional (neither progressive nor regressive). Nevertheless, since income inequality is completely eliminated, a lot of redistribution takes place. Conversely, when the first bracket tax rate is zero, $t_1 = 0$, and the second bracket tax rate is only slightly positive, $t_2 > 0$, the resulting tax schedule will be more progressive than the previous, yet substantially less redistributive. Of course, it is also possible for a tax schedule to be both more progressive and more redistributive.

2.2. Labor supply

Individuals have identical utility functions of the form $g(u)$, where $g(\cdot)$ is strictly increasing and concave. Moreover, we assume that u is quasilinear in consumption, or disposable income, c :

$$u = c - Z(L) \quad (2)$$

where L is the amount of labor supplied, and $Z(\cdot)$ is the cost of labor with $Z(0) = 0$, $Z'(\cdot) > 0$, and $Z''(\cdot) > 0$. For now, we can economize on our presentation of the model by ignoring the role of $g(\cdot)$, which does not affect this part of the analysis. Gross income is $y = wL$. Given the tax structure just specified, individuals face the budget constraint:

$$c_i \leq b + (1 - t_i)y_i \quad i \in \{M, R\} \quad (3)$$

Noting that $L = y/w$, we can write each individual's utility as $u = c - Z(y/w)$. Each individual then chooses y to maximize u subject to her budget constraint.

Carrying through this program, the first-order condition for type- M and type- R individuals is

$$1 - t_i - Z' \left(\frac{y}{w_i} \right) \frac{1}{w_i} = 0$$

which yields the solution:

$$y_i^* \equiv y(t_i, w_i) = Z'^{-1}[(1 - t_i)w_i]w_i$$

This, in turn, produces the following indirect utility function:

$$V_i(b, t_i, w_i) = b + (1 - t_i)y(t_i, w_i) - Z \left(\frac{y(t_i, w_i)}{w_i} \right) \quad i \in \{M, R\} \quad (4)$$

Clearly, since $w_p = 0$, then $y_p = 0$ and indirect utility is:

$$V_p(b) = b \quad (5)$$

We define the compensated elasticity of earned income, y_i , with respect to the net-of-tax rate as:

$$\eta_i \equiv \frac{\partial y_i}{\partial(1-t_i)} \frac{(1-t_i)}{y_i} \quad \text{for } i \in \{M, R\} \quad (6)$$

We assume that this elasticity is nonincreasing in income: $\eta_M \geq \eta_R$.

2.3. Preferences over tax schedules

We define average income as:

$$\bar{y} \equiv \sum_{i \in I} p_i y_i = p_R y_R + p_M y_M \quad (7)$$

On part of the government, we impose a balanced budget requirement, so that tax revenues equal tax expenditures:

$$p_R(t_R y_R - b) + p_M(t_M y_M - b) - p_P b = 0. \quad (8)$$

Solving for b , we obtain:

$$b = p_R t_R y_R + p_M t_M y_M \quad (9)$$

From this we can characterize the preferences over tax schedules of individuals from each group, which we do in the following lemma:

Lemma 1. *Let $\hat{\mathbf{t}}^i \in T$ be each individual's group-specific ideal tax schedule. Then $\hat{\mathbf{t}}^R = (\hat{t}_M, 0)$, $\hat{t}_M > 0$; $\hat{\mathbf{t}}^M = (0, \hat{t}_R)$, $\hat{t}_R > 0$; and $\hat{\mathbf{t}}^P = (\hat{t}_M, \hat{t}_R)$, $\hat{t}_M, \hat{t}_R > 0$. Informally, the progressivity of an individual's ideal tax schedule is nonmonotonic (increasing and then decreasing) in income. In addition, the level of redistribution implied by an individual's ideal tax schedule is decreasing in income.*

Proof. See Appendix A.1.

Critically, this lemma shows that the progressivity of each group's preferred tax schedule is nonmonotonic in income. In particular, the rich prefer a regressive tax schedule, while the middle class and the poor prefer progressive tax schedules. However, the tax schedule preferred by the middle class is more progressive *and* redistributes less than the tax schedule preferred by the poor. It therefore seems *a priori* plausible that tax schedules could emerge that redistribute more, but be less progressive, as the current consensus suggests. We would expect this to be case in systems where political representation enhances the poor's voice.

The reasoning for these preferences is as follows. Because the poor are poorer than the middle class—in fact, the poor have zero earned income in our model—they have little to lose by imposing revenue-maximizing tax rates on all income earners, even those lower in the income distribution, in this case the middle class (t_M). In addition, precisely because they are poor, the marginal cost of this tax is smaller to them than to the middle class. In contrast, the middle class cannot gain from a tax that redistributes between them and the poor, and so prefer to set this rate to zero. Hence, the poor prefer less progressive, but more redistributive taxes than the middle-class. Finally, the rich do not bear any of the cost of the tax rate that applies only to middle-class and poor voters, so they prefer to set this rate at a positive level. In addition, they can only lose by increasing any

tax on themselves, and so prefer to set this rate to zero. The rich therefore prefer regressive tax schedules.

2.4. Preference Aggregation: Political environment

To understand how these preferences translate into redistribution policy we need to describe a model of the political process. Because our problem is bidimensional, we cannot rely on the standard median-voter approach to redistributive taxation: there is no Condorcet winner in our model since any tax-schedule proposal in T can be defeated by a coalition of two groups. To surmount this difficulty, we take the standard approach, building on Lindbeck and Weibull (1987, 1993); Dixit and Londregan (1996), and assume uncertainty about the distribution of voter preferences.

There are two office-motivated parties, A and B , and let $\sigma_{i,j}$ be each individual i 's ideological preference for party $j \in \{A, B\}$. For example, a type- P voter gets utility $g(V_P(\mathbf{t}_A)) + \sigma_{i,A}$ by voting for party A . An individual votes for the party that delivers her the highest utility. Thus, an individual will vote for party A over party B whenever the following condition is satisfied:

$$g(V_i(\mathbf{t}_A)) + \sigma_{i,A} \geq g(V_i(\mathbf{t}_B)) + \sigma_{i,B}$$

The voter in group I who is exactly indifferent between parties A and B has $\bar{\sigma} \equiv \sigma_{i,B} - \sigma_{i,A} = g(V_i(\mathbf{t}_A)) - g(V_i(\mathbf{t}_B))$. Letting $F_i(\cdot)$, $i \in I$ be the cumulative distribution and $f_i(\cdot)$ the probability distribution of $\sigma = \sigma_B - \sigma_A$, then for given policies \mathbf{t}_A and \mathbf{t}_B the fraction of group- I individuals with $\sigma < \bar{\sigma}$, $F_i(\bar{\sigma})$ will vote for party A ; fraction $1 - F_i(\bar{\sigma})$ will vote for party B .

The total votes received by party A can then be written as $\omega_A = \sum_{i \in I} p_i F_i(\bar{\sigma})$. Furthermore, parties A and B face the respective maximization programs:

$$\sum_{i \in I} p_i F_i(\bar{\sigma}(\mathbf{t}_A)) \quad \text{and} \quad \sum_{i \in I} p_i [1 - F_i(\bar{\sigma}(\mathbf{t}_B))]. \quad (10)$$

The final piece in the development of our analytical framework concerns the equilibrium relationship between progressivity and inequality given a probabilistic voting framework. In the Lindbeck and Weibull (1987, 1993) environment, it is straightforward to characterize the equilibrium. First, as is well known, even though the model is ‘‘probabilistic,’’ there is convergence in the platforms proposed by parties because they have no substantive policy preferences and are strictly office seekers. More interestingly for our purposes, under any possible equilibrium *only* progressive tax schedules exist provided that the rich are at least as ideologically polarized as the middle class: $f_M \geq f_R$. With that assumption, and the elasticity assumption made earlier, a necessary and sufficient condition for progressive schedules is the existence of income inequality. The following result captures the core intuition:

Proposition 1. (Progressive Taxation under Democracy) *A symmetric equilibrium tax schedule, $\mathbf{t}^* = \mathbf{t}_A^* = \mathbf{t}_B^*$, exists and is unique. Given assumptions $\eta_M \geq \eta_R$ and $f_M \geq f_R$, progressive tax schedules emerge in equilibrium if and only if there is income inequality.*

Proof. See Appendix A.2.

Moreover, in this setting, progressivity is also enhanced if the labor supply elasticity is decreasing in income and if the rich are more ideologically polarized than the middle class. The intuition is simple: if richer individuals' labor supply is less responsive to taxation and if they are less ideologically polarized, they are more attractive targets for taxation.¹³

3. Analysis and Hypotheses

Making use of this analytical framework, we turn now to analyze the relationship between progressivity and redistribution under democracy. We proceed in two different contexts, one economic, one political. We first study the response in terms of progressivity and redistribution to a change in inequality. Second, we analyze how political representation, and in particular the institutional facilitation of stronger voice by the poor, conditions the relationship.

3.1. Context I: Progressivity and Redistribution given a change in inequality

Proposition 2. (Progressivity and Redistribution) Consider a mean-preserving spread in the income distribution from X to Y such that $p_{X,P} < p_{Y,P}$, $p_{X,M} > p_{Y,M}$, and $p_{X,R} < p_{Y,R}$. Then progressivity and redistribution are higher under (more equal) distribution X than under (less equal) distribution Y .

Proof. See Appendix A.3.

The reasoning for this proposition is as follows: First, the tax rate on rich individuals, t_R falls because the political weight of middle-class voters, who favor a high tax rate on rich voters, diminishes with an increased dispersion of the income distribution. In contrast, the effect of the tax rate on the middle class, t_M , increases. Both rich and poor voters prefer a higher t_M than middle-class voters, which tends to increase t_M relative to the lower inequality state. Thus, as inequality increases, the tax schedule becomes clearly less progressive.

With an increase in t_M redistribution decreases for several reasons. First, an increase in t_M increases the tax burden on middle-class voters, which reduces redistribution. Second, as the proportion of middle-class incomes declines and the proportion of poor incomes increases, revenue raised by this tax falls. Finally, the fall in revenue from t_R is unlikely to be replaced by the increased revenue from t_M , both because the labor-supply elasticity of middle-class workers is at least as high as for rich workers, and simply because the rich have higher incomes than either.

The central implication from this simple exercise is clear: *ceteris paribus*, as progressivity decreases, so does redistribution, a change that is in turn associated with an increase in the level of post-tax, post-transfer inequality. This result calls into question the notion, widely shared and often discussed, that taxes do little of the redistributive work in fiscal systems. Leaving aside the fact that in fiscal incidence terms such a statement is a bit of an oxymoron, this result provides important correction to a dominant view in the political economy of redistribution. Our model suggests that

¹³This result demonstrates most clearly the value to modeling labor supply on the extensive margin.

for any given level of effort and the progressivity of benefits, a change in the progressivity of taxes does have a significant and positive effect on redistribution.

This result also raises a question for another stream of research, namely the relationship between progressivity and the overall scope of redistribution in society. Recall that, as discussed above, several previous contributions point to the existence of a trade-off between the two, especially in those polities where the system of representation has facilitated the creation and sustainability of pro-redistributive cross-class coalitions (Kato 2003; Beramendi and Rueda 2007). Does the fact that progressivity exerts a positive marginal effect on redistribution imply that the relationship between redistribution and the overall tax structure is, contrary to what was previously held, also linear and positive? To address this question we need to evaluate the relationship between progressivity and redistribution under institutional conditions that facilitate the political voice of low income citizens, and therefore an increase in the overall *level* of redistribution.

3.2. Context II: Political Representation

The political influence by competing income groups works through many potential channels, some outside the democratic process, such as lobbying, others via the system of representation itself (Przeworski 2010). Here we focus on the latter. To account for the presence of political institutions, we add an additional, legislative-bargaining stage to the political problem and assume the existence of three, class-representative political parties. Political parties are policy motivated and the preferences of each party reflect those of group whom they represent, R , M , or P . Second, to keep the analysis simple, we assume that voters make their choices sincerely (in the two-party case above, voter's choices can be either sincere or strategic). Third, the party with the largest share of votes is chosen with certainty as the formateur. The formateur then chooses a coalition partner (to maximize its policy preferences) and the parties set policy by engaging in Nash bargaining. Finally, we assume that ideological preferences are distributed symmetrically around the neutral (i.e., $\sigma = 0$) voter.

Like other models that compare majoritarian and proportional representation (see, e.g., Austen-Smith (2000) and Iversen and Soskice (2006)), our model analyzes politics in majoritarian institutions as two-party competition (as in Duverger's Law) with each party vying for the decision votes of middle-class voters. In this case, policy will always reflect middle class interests.¹⁴ In contrast, under proportional representation, there are three parties, each representing a different economic constituency: P, M, R . Furthermore, policy making under proportional representation features coalition government, inasmuch as the winners do not take all. Therefore, following an election stage, we adopt a legislative bargaining stage wherein parties bargain with other parties in coalitions. Using this political environment, our analysis yields the following proposition.

Proposition 3. (Political Representation)

¹⁴We simplify their approach somewhat by assuming that there is never a danger of either party deviating to a rich or poor person's preferred platform. Dropping this assumption has uncertain effects. It certainly does not change the result that PR systems redistribute more than majoritarian. The effects of dropping the assumption on progressivity are less clear. Crucial in this regard is their non-regressivity assumption (Iversen and Soskice 2006: 167), an assumption we do not make in this paper.

Suppose that under majoritarian representation, tax policy is coincident with middle-class preferences: $\mathbf{t}^* = \hat{\mathbf{t}}_M$. Then redistribution is higher and progressivity is lower under proportional representation than under majoritarian representation.

Proof. See Appendix A.4.

As we have previously demonstrated in Lemma 1, under simple assumptions, individuals' ideal tax schedules are nonmonotonic (increasing and then decreasing) in income. In contrast, the amount of redistribution that occurs under these ideal tax schedules is decreasing in income. From this it is a short step to the conclusion that, all else equal, majoritarianism increases progressivity but reduces redistribution relative to proportional representation. The translation of these preferences into policy differs by system of representation.

Parties in majoritarian systems give greater weight to middle-class voters, while proportional systems give greater weight to poor voters. This effect may reflect that the key decision maker is the median legislator left of median (Austen-Smith 2000) or the formation of coalitions where parties representing the poor and the middle classes must compromise with each other (Iversen and Soskice 2006). In either case, given that middle-class voters prefer more progressivity and less redistribution than poor voters, tax schedules will be more progressive but will redistribute less under majoritarian representation. In contrast, to the extent that the system of representation translates the poor's preferences into policy¹⁵, such systems will exhibit less progressive tax schedules but more redistribution.

This results suggests that as the political influence of the poor increases, through this and plausibly other channels not modelled directly in this paper, increases in the overall size of redistribution (as supported by low income voters) will require partial sacrifices in terms of the overall progressivity of the income tax design. The higher levels of redistribution are viable because the "sacrifice" in terms of design allows for a much larger revenue base, along the lines of the non-monotonic preference scheme above. In addition, such a design allows the tax burden on the middle classes to be reduced while at the same time facilitating the provision of encompassing systems from which they also benefit.

Formally, the "sacrifice" is best captured through the relationship between the two components of the tax function: the progressivity of the tax design and the level of flat rate tax income. Proposition A.4 implies that as the former increases the latter must decrease for the mapping of pre-tax to post-tax income, that is for the overall level of redistribution, to remain the same. In other words, there is a ceiling to the amount of redistribution that can be achieved via tax progressivity. Paradoxically, if political actors want to maximize redistribution, as it plausibly happens in systems where the poor exercise significant political influence, the optimal combination is not one that keeps progressivity at its maximum. Rather, it is one in which the flat rate is marginally higher and is kept at intermediate levels.

This points to a more general theoretical conclusion: across the distribution of fiscal systems, there is an inverse relationship between their progressive component (more prominent when

¹⁵We are aware of two important implicit premises in our model. First, there is no second dimension (nation, race, or other forms of heterogeneity) driving the poor away from their predicated material interest (Shayo 2009; Rueda and Stegmüller 2019).

the overall level of redistribution is lower) and their proportional one (more prominent when the overall level of redistribution is higher). In addition, an important corollary about the role of political institutions also follows: in systems where low income voters are influential the ratio between the two components of the tax function (progressive to flat tax rate) will be smaller.

3.3. Summary of Empirical Implications

The theoretical analysis above yields empirically testable implications about the relationship between progressivity and redistribution:

1. *On the marginal effect of tax progressivity on redistribution (Proposition 2):* For any given level of effort and the progressivity of benefits, a change in the progressivity of taxes does have a significant and positive effect on redistribution (H1)
2. *On the relationship between redistribution and tax designs (Proposition 3):* There is a negative association between the progressivity of the tax system and the level of proportional (flat rate) taxation (H2).
3. *Corollary:* In systems of representation that allow for higher political influence by low income groups (PR vs SMD), the ratio of progressivity to proportional taxation decreases.

4. Data and Empirical Strategy

We turn now to the empirical assessment of these implications, more specifically, we test hypotheses H1 and H2, and then provide an empirical illustration of our corollary. We begin by describing a novel measurement approach specifically designed to exclude changes in market distributions induced by the very fiscal tools under analysis. On the basis of these analyses we recover our own estimates of λ and τ for 21 OECD countries. The limiting factor in the selection of our sample is the availability of high-quality measures of redistribution (described below). Thus our analysis sample consists of 203 country-years. The countries included in our analysis are Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Japan, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, United Kingdom, and the United States. Table B.1 in the appendix shows which years we observe for each country. The median number of years available by country is 11 (17 countries provide at least 8 years of observations). Thus, we have enough information to employ within-country designs in our statistical analyses.

4.1. Measuring tax progressivity

We measure tax progressivity using a policy simulation approach. This choice is driven by our desire to capture the intended effect of tax policy set by governments given an existing distribution of market incomes. In other words, we want to measure the ‘pure’ structure of tax policy *net* of possible changes in the distribution of market incomes. Excluding changes in market distributions from the calculation of policy measures is important. As Beramendi (2001) notes most, if not all,

policies create second order effects. Policies, when implemented (or announced), do not simply shift the equilibrium distribution of post-government economic outcomes. They also alter the distribution of market outcomes, since individuals react to policy changes. Thus, policies “generate behavioral responses from market actors, who adjust their economic decisions to the nature and changes of policy interventions” (Beramendi 2001: 5).¹⁶ Respondents’ market and disposable income derived from household surveys are a mixture of “first order” effects of welfare policies and “second order” effects of individuals’ behavioral responses (Boadway and Keen 2000: 760). Using policy simulation allows us to separate these two effects. We use the OECD tax-benefit simulation model, which contains detailed rules of both tax and benefit policies covering OECD countries between 2001 and 2015 (OECD 2007). It allows us to calculate the income consequences of each country’s tax and benefit structure for households at different point in the income distribution. We then summarize these income effects using progressivity measures, described below.

Simulation details In order to capture a wide range of realistic individual circumstances, our simulation includes four family types: single individuals, single parents with children, and married couples with and without children. In families there is one active working-age individual. Our policy effect simulation computes the amount of taxes (and benefits) across the earnings spectrum for each family type. To normalize difference in average earnings between countries, all earnings are expressed in percentiles relative to those of the average production worker (APW) in a country. They range from 50 to 200 percent of the APW wage.¹⁷ Thus, our results capture the effects of taxes on the incomes of working-age individuals and their families. The taxes included in the simulation are Personal Income taxes, as well as Employee and Employer paid Social Security Contributions. This calculation yields a data set of over 250,000 records of pre and post tax incomes (each country \times family type \times APW income level).

Tax progressivity With these data on the income consequences of taxes at hand, we can calculate our measure of tax progressivity. We use a tax function approach to approximate each country’s tax system, following a long tradition in public finance, where this class of tax-transfer models has been introduced by Feldstein (1969), and later extended to dynamic models by Persson (1983) and Benabou (2002).

In this approach, tax revenue is a function of individual income w_i , since income is the only factor on which taxes can be conditioned on. The retention function of an individual’s income is given by:¹⁸

$$T(w_i) = w_i - \lambda w_i^{1-\tau}. \quad (11)$$

¹⁶This issue is widely acknowledged in macro-economics, where simulating the effect of policy interventions requires not only a knowledge of existing and future policy rules, but also the knowledge of individual preference parameters to predict individuals’ behavior responses at both extensive and intensive margins (for extended discussions of issues surrounding policy effect analysis (including shortcomings of our current static approach) see, e.g., Atkinson, King, and Sutherland 1983; Feldstein 1983; Figari, Paulus, and Sutherland 2015).

¹⁷This is a synthetic (not empirical) income distribution anchored at the wage of the average production worker. It thus holds constant factors such as income-specific incidence of unemployment spells during business cycles, or country-specific gender disparities in occupations.

¹⁸The marginal tax rate on an individual’s income is thus given by $T'(y) = 1 - \lambda(1 - \tau)y^{-\tau}$

It implies a non-linear mapping between after tax earnings x_i and pre-tax earnings w_i :

$$x_i = \lambda w_i^{1-\tau}, \quad (12)$$

where $1 - \tau$ is the elasticity of before-tax to after-tax income. Thus, τ is a straightforward measure of the progressivity of the tax system. Generally, a tax system can be defined as progressive if the ratio of the marginal tax rate to the average tax rate is greater than one for a given level of income (e.g., Slitor 1948: 310). In terms of our tax function this entails:

$$\frac{T'(w_i)}{T(w_i)/w_i} = \frac{1 - \lambda(1 - \tau)w_i^{-\tau}}{1 - \lambda w_i^{-\tau}} \quad (13)$$

This expression is larger than 1 when $\tau > 0$ yielding a progressive tax structure. If instead $\tau < 0$ the ratio of marginal to average tax is less than one at a given level of income yielding a regressive tax structure. The parameter λ shifts the tax function for a given level of τ and we interpret it as a parameter capturing the ‘proportional’ component of the tax system.

We fit the tax function in (12) to the pre-post income data from our tax simulations, and obtain estimates for λ and τ for each country and year between 2001 and 2015.¹⁹ Table B.2 in the appendix provides an overview of λ and τ estimates.

4.2. Definition of other variables

Redistribution. We use data from the OECD income distribution database, which compiles income information from administrative sources and household panel surveys in advanced industrialized countries (e.g., Förster and Pearson 2002).²⁰ The data sources used are of high quality, and estimates derived from it are comparable to those from the Luxembourg Income Study (OECD 2012). The advantage of using this data source is that it provides us with wider data coverage in terms of available country-years. For each country and year we have information on the Gini index of inequality of (square-root scale) equivalized household incomes, (i) at market incomes (before taxes and transfers), and (ii) at disposable incomes (after taxes and transfers). In line with existing studies (e.g., Kenworthy and Pontusson 2005: 455), we operationalize the extent of redistribution as the difference between (i) and (ii).²¹

Benefit progressivity Our measure of benefit progressivity also relies on data from our simulation model. As in our policy simulation for tax structures, we calculate benefits for four different types of households at varying points in the income distribution (ranging from 50 to 200 percent of the average production worker’s income). When calculating the income effect of benefits, we include Unemployment, Social Assistance, Housing, Family, and in-work benefits. Based on this detailed

¹⁹The model approximates observed income patterns quite well with R^2 values usually above 0.99.

²⁰The data source for many countries is the EU Survey of Income and Living Conditions, which is one of the primary sources of European Union policy making. Other countries rely on high quality national household panels, their census, or register data.

²¹An argument against this “absolute” definition of redistribution is that it does not take into account initial levels of inequality (as “relative” measures, which express the difference in percentage of initial levels, do). But note that in our empirical specifications we estimate models that rely only on *changes* in redistribution, as well as a model which explicitly includes previous levels of redistribution.

data set we calculate the Kakwani progressivity measure of the distribution of benefits (Kakwani 1977b; Kakwani and Podder 1976). See Beramendi and Rehm (2016) for more discussion of this measure.

Social spending and proportion of elderly. Two key variables we adjust for in our analyses are existing levels of spending on social programs (e.g., International Labor Organization 2008: 130) and a country’s (changing) share of the elderly population. We measure social spending as total public expenditure in percent of gross-domestic product, available in the OECD’s SOCX database and the elderly population as the share of individuals aged 65 and older (also obtained from the OECD).

4.3. Statistical specifications

We estimate several empirical specifications to study the relationship between tax progressivity and redistribution after accounting for levels of social spending and progressivity of benefits. We rely on two-way fixed effects models, which rely on within-country changes in all variables while adjusting for common time shocks. More extended specifications explicitly account for autoregressive structures, pre-determined regressors, non-unit-constant shocks and slope heterogeneity. Before describing our main model specifications, note that our key measures are sufficiently distinct to jointly include them in a model analyzing redistribution. The average correlation between benefit progressivity and tax progressivity is 0.36. This correlation decreased from 0.43 in the 2000s to 0.27 in the 2010s. The correlation of benefit and tax progressivity with spending is -0.05 and 0.03 , respectively. All variables also show significant within-country changes over time.

Our first empirical specification for redistribution in country i in year t , y_{it} , is given by

$$y_{it} = \alpha\tau_{it} + x'_{it}\beta + \phi_i + \zeta_t + \epsilon_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T_i. \quad (14)$$

Here, redistribution is a function of tax progressivity τ_{it} with associated coefficient α and vector of controls (including benefit progressivity, social spending, and share of elderly population). We include country fixed effects ϕ_i , which capture (time-constant) unobservables on the country level as well as time fixed effects ζ_t (Baltagi 2013: 39f.). While this setup is the ‘default’ in many empirical analysis of panel data, it imposes the restriction that all units are affected by time-specific shocks in the same way, a point to which we return to below. Our second empirical specification accounts explicitly for the fact that residuals for the same country are likely correlated over time. We estimate a fixed effects specification where the residuals follow a stationary AR(1) process (taking into account the unbalanced nature of our dataset; cf. Baltagi and Wu 1999):

$$y_{it} = \alpha\tau_{it} + x'_{it}\beta + \phi_i + \epsilon_{it}, \quad \epsilon_{it} = \rho\epsilon_{it-1} + v_{it} \text{ with } |\rho| < 1 \quad (15)$$

In our third, more involved, specification we estimate a dynamic panel specification by including lagged values of redistribution, $y_{i,t-1}$, in our fixed effects model:

$$y_{it} = \rho y_{i,t-1} + \alpha\tau_{it} + x'_{it}\beta + \phi_i + \zeta_t + \epsilon_{it} \quad (16)$$

Simply introducing a lagged dependent variable into models with country fixed effects introduces bias (the lagged dependent variable violates strict exogeneity) particularly in a small-T context (Nickell 1981). We estimate the model using a difference GMM (generalized method of moments) estimator suggested by Arellano and Bond (1991). The estimator operates on a first differenced system of equations, where instruments are generated (i) for the lagged dependent variable from its own lags in each time period, (ii) for the first difference equation from differences of exogenous covariates (cf. Arellano and Bond 1991; Holtz, Newey, and Rosen 1988).²² This specification removes two periods and reduces our sample size to 141 country-years.

5. Results

5.1. Progressivity and redistribution

Table I shows estimates from six specifications. We begin with our basic two-way fixed effects model and sequentially enter levels of social spending, benefit and tax progressivity. For ease of interpretation we standardize all inputs to have mean zero and unit standard deviation. Our results in the first column of Table I show that, not surprisingly, higher levels of spending achieve more redistribution. Of more interest is the role of the progressivity of the tax and transfer system. Interestingly, column (2) shows that changes in the progressivity of benefits have no clear relation to increases or decreases in redistribution. In contrast, it is tax progressivity that has a clear, statistically significant, impact. A standard deviation increase in τ increases redistribution by 0.4 (± 0.1) standard deviations. This finding is also illustrated in Figure I, where we plot expected values of redistribution at varying (unstandardized) levels of benefit and tax progressivity. A change from $\tau = 0.15$ (just below the median) to the 75% percentile ($\tau = 0.20$) increases redistribution by 1.7 points, or 15 percent. The corresponding change for benefit progressivity is negligible.

In columns (4) of Table I we allow for serial correlation of residuals (while taking into account unevenly spaced observations in each country) as specified in equation (15). The estimated correlation ($\rho = 0.6$) signifies that residuals are indeed strongly serially correlated. Accounting for this correlation also reduces the regressor fixed-effects dependence in the model.²³ Under this specification we find the coefficient for spending significantly reduced, while there is little change for the estimated impact of changes in benefit progressivity. The effect of tax progressivity is reduced by a third, but still of substantive magnitude and clearly statistically different from zero. A standard deviation increase in progressivity is associated with a 0.28 (± 0.07) standard deviation increase in inequality reduction.

²²We did not make this model our preferred specification since it depends much more on specific modeling choices (such as the depth of lags to include) impacting the validity of the created instruments. We have limited the lag depth to 8 in our analysis, but ensured that unlimited depth does not change our results. To inspect the appropriateness of the model, we conduct two tests. One key issue is the assumption that the time-dependence follows a first-order Markov process (so that taking second-order lags creates instruments). This assumption can be tested by inspecting the auto-correlation of residuals after the model is estimated (ensuring that the first differenced residuals do not exhibit second-order autocorrelation). We also use the overidentifying restrictions to test the joint validity of the moment conditions of the GMM estimator using the Sargan-Hansen test.

²³The correlation between regressors and country fixed effects in a fixed effects specification is -0.293 , which is reduced to -0.059 in specification (4).

Table I.
Redistribution as function of spending and progressivity.

	(1)	(2)	(3)	(4)	(5)	(6)
Spending levels	0.844 (0.119)	0.842 (0.120)	0.989 (0.117)	0.501 (0.070)	0.476 (0.089)	0.343 (0.102)
Benefit progressivity		−0.036 (0.231)	−0.092 (0.196)	−0.071 (0.059)	−0.131 (0.081)	−0.054 (0.093)
Tax progressivity			0.439 (0.098)	0.284 (0.072)	0.243 (0.091)	0.229 (0.089)
ρ				0.607	0.534 [‡]	0.583 [‡]
Two-way fixed effects	✓	✓	✓	✓	✓	✓
Δ economic vars.						✓
Estimator	FE	FE	FE	FE AR(1)	GMM	GMM
R-squared [†]	0.31	0.32	0.44	0.53	0.42	0.54
N	203	203	203	182	141	141

Note: Unbalanced panel of 21 OECD countries, 2001–2015. All inputs normalized to mean zero and unit standard deviation. Cluster-robust standard errors.

Specifications: (1)-(3) Two-way fixed effects models (country and year). Average T=10.7. (4) AR(1) model with fixed effects (Baltagi and Wu 1999). (5) LDV model with fixed effects (Arellano and Bond 1991), GMM IV estimates; estimated on differenced system, using lagged LDV and differenced covariates as instruments: AR(2) test of residuals $p = 0.652$. Specification (6) is (5) with added economic variables (first differences in inflation, real GDP growth, and unemployment rate). AR(2) test $p = 0.157$. All models include the share of the 65+ population.

[†] Refers to “within-panel” R-squared (calculated using doubly demeaned data).

[‡] Coefficient on lagged dependent variable.

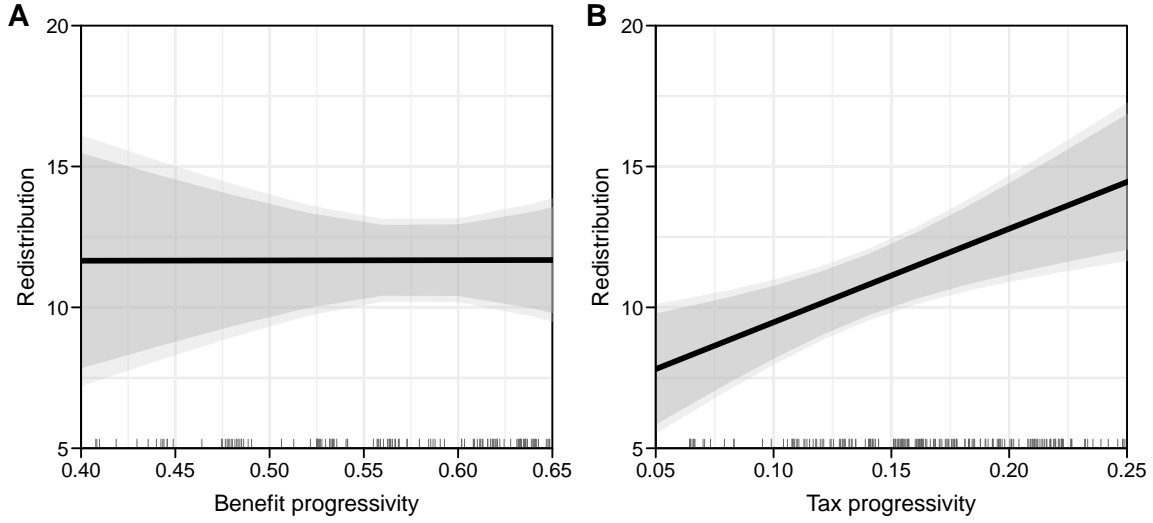


Figure I.
The structure of taxes and transfers and redistribution

Expected value of redistribution at levels of tax progressivity and benefit progressivity. Shaded areas represent 90 and 95% confidence intervals. Based on two-way fixed effects model fitted to panel of 21 OECD countries, 2001-2015.

We estimate a dynamic panel model in column (5) following the specification in equation (16). As expected, we find strong persistence of patterns of inequality reduction. The estimate for ρ , the coefficient for y_{it-1} , is sizable and statistically different from zero. Thus, redistribution in year t is in large parts determined by the amount of redistribution carried out in year $t - 1$. In this setting, what is the effect of a change in the progressivity of the tax and transfer system? Even in this much more involved specification we find clear evidence for the substantive and statistical relevance of the progressivity of a country's tax structure. The contemporary effect of a unit-change in tax progressivity on redistribution is 0.24 standard deviations, while the long run effect (taking into account both the contemporary change and its feedback via lagged redistribution) is 0.55 (s.e.= 0.22). Finally, this is also confirmed in specification (6) where we add variables representing economic conditions that might effect achieved redistribution in a mechanical way, namely changes in inflation, real GDP growth and unemployment.

5.2. Specification tests

Before we proceed to a discussion of the political significance of our findings, we subject our results to a number of specification tests. We start with a model that allows for heterogeneous common shocks and flexible cross-sectional dependence. As discussed above, the traditional two-way fixed effects model setup assumes that country and time effects enter the model additively. Relaxing this assumption can be achieved by specifying a model with interactive fixed effects (Bai 2009) implemented via a factor structure:

$$y_{it} = \alpha\tau_{it} + x'_{it}\beta + \xi'_i F_t + \epsilon_{it}. \quad (17)$$

Table II.
Specification tests. Estimate of tax progressivity (standard errors in parentheses).

	Tax progressivity
(1) Interactive fixed effects estimator	
One common factor (r=1)	0.411 (0.084)
Two common factors (r=2)	0.438 (0.114)
(2) Heterogeneous panel estimator (MG)	0.613 (0.293)
(3) Bayesian TSCS model with two-way RE	0.420 (0.068)
(4) Balanced panel (multiple imputation)	0.397 (0.133)
(5) Percentile-t wild bootstrap imposing null	$p=0.018$

Specifications: (1) Interactive fixed effects estimator with 1 and 2 common factors (Bai 2009). (2) Allows for heterogeneous regressor slopes and time trends via Pooled Mean Group estimator (Pesaran and Smith 1995). (3) Bayesian hierarchical model with country and year random effects, regressor RE dependence via Mundlak device. Based on 20,000 MCMC samples. (4) Balanced panel, N=313. Regression imputation using country-specific time trend (M=100). MI corrected standard robust errors. (5) Bootstrapped, two-way clustered standard errors. Entry is test of significance using 1,000 percentile-t wild bootstrap samples imposing the null and clustering on both country and time. (Cameron, Gelbach, and Miller 2008).

Here, F_t is a vector of r common factors, which represents the structure of unobservables shared by cross-sectional units in a given year, and ξ_i is a $r \times 1$ vector of corresponding factor loadings. ϵ_{it} are idiosyncratic errors. To see how this generates heterogeneity, think of F_t as a vector of common shocks (e.g, the financial crisis). Then ξ_i represents the heterogeneous impact of this shock on country i . We treat both ξ_i and F_t as fixed-effects parameters to be estimated using principal component methods (cf. Bai 2009: 1236f.).²⁴ The model makes no assumption about the mean of F_t or the structure of its over time dependency (see Pesaran 2006; Bai 2009 for a discussion of its asymptotic properties). Furthermore, we allow spending, benefit and tax progressivity to evolve endogenously:

$$x_{it} = \mu_i + \theta_t + \sum_{k=1}^r a_k \xi_{ik} + \sum_{k=1}^r b_k F_{kt} + \sum_{k=1}^r c_k \xi_{ik} F_{kt} + \pi_i' G_t + \eta_{it} \quad (18)$$

Here, a_k , b_k , and c_k are scalar constants and G_t is a different set of common factors (that do not enter in the outcome equation). Thus, covariates x_{it} can be correlated with ξ_i alone, or with F_t alone, or simultaneously with ξ_i and F_t . Specification (1) of Table II show that a model with interactive fixed effect estimated with one common factor produces estimates for the impact of τ that are very close to our main results. Extending the model to two common factors increases the impact of τ : a standard deviation increase in tax progressivity leads to an increase in redistribution of 0.44 (± 0.12) standard deviations.

²⁴Note that we have to choose the number of factors a priori. One factor is enough to introduce cross-sectional dependence (Pesaran 2006: 972) and allow for interactive fixed effects. In our empirical implementation below we also estimate a model with two factors (which is still supported by the data).

So far we have assumed that the coefficients on all covariates are constant over countries. If there is heterogeneity in some slopes, e.g., if the impact of spending varies over countries, does it change our understanding of the (average) effect of tax progressivity? In specification (2) we employ an estimation strategy that explicitly allows for heterogeneity in slopes (including country-specific time trends) employing the mean group estimator of (Pesaran and Smith 1995).²⁵ Our results in specification (2) show that this has little impact on the average effect of tax progressivity, but leads to increased standard errors (representing the increased heterogeneity in the model). That notwithstanding, our core result on the role of tax progressivity is confirmed.

The final three specifications are more technical in nature. In (3) we estimate our model in a Bayesian framework providing partial pooling estimates for both country and time random intercepts. See Shor et al. (2007) for the advantages of Bayesian inference with TSCS data. We account for regressor random effect dependence in both dimensions using the Mundlak specification (Rendon 2012).²⁶ We find little change in our substantive results. In specification (4) we create a balanced panel (with 313 country-years) by filling in values for redistribution under a MAR assumption. Missing years are primarily the results of lack of household panel data in the OECD Income Distribution database, making it more likely that the missingness process is not MNAR. Note that we have complete information on all years for our measures of progressivity, as well as for social spending and the share of the elderly. We create 100 imputed data sets and adjust our standard errors for the increase in imputation variance. Our results show little substantive change. Finally, specification (5) we test the statistical significance of the estimate for τ using a percentile-t wild bootstrap imposing the null hypothesis and taking into account clustering on both countries and years. See Cameron, Gelbach, and Miller (2008) for an extended discussion. The entry in Table II is the p-value from a two-sided test of the resulting empirical distribution.

²⁵The central idea is to estimate N group-specific coefficients for all covariates and a time trend using OLS and then to combine the estimated coefficients across groups. Since this strategy relies on within-country estimates we only include countries with longer series (this amounts to excluding Australia, Japan, New Zealand, and Switzerland). Nonetheless, the pooled mean group estimator assumes large samples in *both* dimensions (N and T) and thus our specification should be seen as providing only suggestive evidence on heterogeneity.

²⁶We choose non-informative priors with mean zero and variance 100 for all regression type parameters. Variance priors are inverse gamma with shape and scale parameters set to 0.005.

5.3. Progressivity and flat-rate tax levels

We now turn to an empirical investigation of H2, which states that we should expect a negative relationship between tax progressivity and the flat-rate tax parameter $1 - \lambda$. We model the within-country relationship between changes in τ and changes in $1 - \lambda$. To ease interpretation, we standardize both inputs and outputs to mean zero and unit standard deviation. We include both linear and quadratic terms of the tax progressivity parameter, because the inspection of a LOESS smoother suggests that a quadratic approximation captures the relationship between both variables rather well (cf. Figure B.2). But note that our substantive results also obtain when using a linear term only.

Table III.
Within-country relationship between λ and progressivity τ .

	(1)	(2)	(3)	(4)
$\partial\lambda/\partial\tau$	-0.734 (0.054)	-0.672 (0.072)	-0.877 (0.056)	-0.859 (0.023)
Two-way fixed effects	✓	✓	✓	✓
Country-time trends		✓	✓	—
Estimator	FE	FE	GMM	IFE
R-squared [†]	0.64	0.73	0.93	0.77
N	315	315	273	315

Note: Balanced panel of 21 OECD countries, 2001–2015. Both τ and λ normalized to within-country mean zero and unit standard deviation. Estimated equation is quadratic in τ ; table entry is marginal effect of τ with cluster-robust standard errors.

Specifications: (1) Two-way fixed effects models (country and year). Specification (2) adds country-specific linear time trends. (3) LDV model with fixed effects (Arellano and Bond 1991), GMM IV estimates; estimated on differenced system, using differenced covariates and lagged LDVs as instruments. Number of included instruments reduced using principal components (28 components with eigenvalue > 1 , $R^2 = 0.81$). AR(2) test of residuals $p = 0.156$. (4) Interactive fixed effects model (Bai 2009) with 2 common factors.

[†] Refers to “within-panel” R-squared (calculated using doubly demeaned data).

Table III shows marginal effect estimates of tax progressivity on flat rate taxes. Like in Table I, we start with a two-way fixed effects model in specification (1) followed by the inclusion of country-specific time trends. We find that a standard deviation increase in tax progressivity decreases flat-rate taxes by about $0.7(\pm 0.07)$ standard deviations. In specification (3) we estimate a dynamic panel specification (similar to specification (5) in Table I) allowing for state-dependence of flat-rate taxes. Finally, in specification (4) we estimate an interactive fixed effects model with two common factors as discussed in section 5.2. Both specifications, despite differing fundamentally in their assumptions point to a very similar impact of tax progressivity: a unit change in progressivity of the system is related to a ≈ 0.8 standard deviation decrease in the flat-rate tax parameter. This provides empirical support for the relationship derived from our model and expressed in Hypothesis 2.

This is an important finding that sheds light on the ambiguity found in previous contributions on the notion of “redistribution within one class” (Lindert 2004; Cusack and Beramendi 2006),

that is on the idea that for sufficiently high levels of redistribution to be politically feasible workers and consumers must foot a significant share of the bill. Our model zooms into the mechanism behind this pattern: as progressivity increases, the size of the tax base on which revenues are collected declines, eventually yielding a relatively smaller pool of revenues to be shared. Our findings display a pattern that is consistent with this logic, but are now on the basis of new, and more precise, measurement strategy and more rigorous empirical tests.

6. Political Institutions and the Progressivity-Redistribution Link

We turn now to assess the corollary of our model regarding the role of different systems of representation. Our theoretical argument suggests that as the political influence of the poor increases, the ratio of progressivity to proportional taxation decreases. If sufficiently high redistribution comes at the expense of a partial sacrifice of progressivity, it should follow that in democracies where political institutions (PR) facilitate stronger political influence by the poor, the ratio between τ and $1 - \lambda$ is lower.

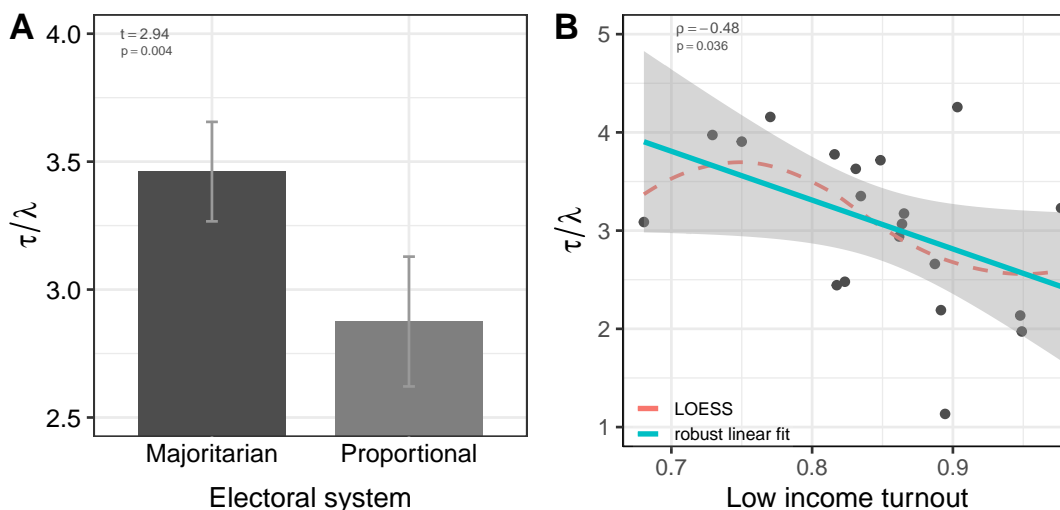


Figure II.
Electoral influence of the poor and proportional taxation

A shows the average value of τ/λ in majoritarian and proportional electoral systems. Error bars represent 95% confidence intervals. The difference in means of τ/λ between electoral systems is 0.59 ($t=2.9$, $p=0.004$). B plots τ/λ against the turnout rate among low income citizens estimates from the CSES (cf. appendix C). Superimposed is a robust linear model fit with 95% confidence bands and a LOESS smoother (with span 1; dashed line). The Spearman correlation between the two measures is -0.48 (exact $p=0.036$).

To address this corollary directly, panel (A) plots the average (and 95% confidence intervals) of the ratio of progressivity to λ in majoritarian versus PR electoral systems.²⁷ In line with our

²⁷We exclude mixed electoral systems in this calculation. However, note that including them in the reference group does not substantively alter our finding.

theoretical expectations, in countries with majoritarian electoral systems we find the ratio to be 0.59 (± 0.20) units greater than in countries relying on proportional representation.

In Panel (B) we use a different measure of political influence: the turnout rate among low-income citizens. Using data from the Comparative Study of Electoral Systems from 2001 to 2016, we calculate turnout rates in the bottom two quintiles for each country in our sample. See appendix C for details on measurement and data harmonization, and how we deal with survey missing responses. Panel (B) plots low-income turnout against the τ/λ ratio and superimposes a robust linear model (with 95% confidence bands) and a LOESS smoother. In line with the result in panel (A), it shows that increasing turnout among lower income citizens decreases the ratio between τ and λ .

7. Conclusion

What governs the relationship between progressive taxation and redistribution? A layman's view would suggest that both are one and the same. And yet, the dominant view so far seems to contend that effective redistribution requires a sacrifice in terms of the progressivity of the design of tax structures. Accompanying these analyses, mostly based on cross-national macro-level data, the case is often made that redistribution is something driven by the "spending side" of the fiscal system, adding yet another layer of ambiguity to the problem.

This paper has developed a formal analysis of the political underpinnings of progressive taxation and a new, incidence-free, measurement strategy to revisit and clarify this relationship. Assuming that actors maximize their preferred level of redistribution, we show formally that: (1) there is a negative relationship between inequality and tax progressivity, and a positive relationship between tax progressivity and redistribution; (2) there is a non-monotonic relationship between income and preferences for progressive taxation: the poor actually may prefer partial sacrifices of progressivity to secure a larger pool of revenue from which to benefit. This larger pool of revenue would come through a larger flat-tax rate. This suggests two empirical expectations: (1) there is a positive marginal impact of tax progressivity on redistribution, contrary to the notion that all the action is on the same time; (2) at the same time, as the political influence of the poor increases, the relative balance between progressive vs flat-rate taxes changes in favor of the latter. Our empirical analyses lend empirical support to both contentions, thus reconciling the layman's view with the need to pay attention to both the marginal effect of progressive taxes and the size of the revenue pool when the role of income taxation in redistributive politics.

Our findings also suggest several paths for future research efforts. The most obvious one concerns the systematic exploration of the non-linear combination of tax progressivity and fiscal effort in a multivariate context. Our multivariate models have uncovered a robust linear effect of progressive taxation. At the same time, our politico-economic analysis of the institutional underpinnings of these relationships suggest that for a subset of countries (PR) the relationship changes. Exploring this contention in a rigorous, systematic way is the focus of ongoing research efforts.

References

- Alt, J. E. 1983. "The Evolution of Tax Structures." *Public Choice* 41 (1): 181–222.
- Alvaredo, F., L. Chancel, T. Piketty, E. Saez, and G. Zucman. 2018. *World inequality report 2018*. Belknap Press.
- Arellano, M., and S. Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies* 58 (2): 277–297.
- Atkinson, A. B., and J. E. Stiglitz. 2015. *Lectures on Public Economics: Updated Edition*. Princeton University Press.
- Atkinson, A. B., T. Piketty, and E. Saez. 2011. "Top incomes in the long run of history." *Journal of Economic Literature* 49 (1): 3–71.
- Atkinson, A., M. King, and H. Sutherland. 1983. "The Analysis of Personal Taxation and Social Security." *National Institute Economic Review* 106 (1): 63–74.
- Austen-Smith, D. 2000. "Redistributing income under proportional representation." *Journal of Political Economy* 108 (6): 1235–1269.
- Bai, J. 2009. "Panel data models with interactive fixed effects." *Econometrica* 77 (4): 1229–1279.
- Baltagi, B. H. 2013. *Econometric Analysis of Panel Data*. New York: Wiley.
- Baltagi, B. H., and P. X. Wu. 1999. "Unequally spaced panel data regressions with AR (1) disturbances." *Econometric Theory* 15 (6): 814–823.
- Benabou, R. 2002. "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?" *Econometrica* 70 (2): 481–517.
- Beramendi, P. 2001. "The Politics Of Income Inequality In The Oecd: The Role Of Second Order Effects." Luxembourg Income Study Working Paper No. 284.
- Beramendi, P., and D. Rueda. 2007. "Social democracy constrained: indirect taxation in industrialized democracies." *British Journal of Political Science* 37 (04): 619–641.
- Beramendi, P., and P. Rehm. 2016. "Who Gives, Who Gains? Progressivity and Preferences." *Comparative Political Studies* 49: 529–563.
- Binmore, K., A. Rubinstein, and A. Wolinsky. 1986. "The Nash bargaining solution in economic modelling." *The RAND Journal of Economics*: 176–188.
- Boadway, R., and M. Keen. 2000. "Redistribution." In *Handbook of Income Distribution*, ed. A. B. Atkinson and F. Bourguignon. Vol. 1. Elsevier.
- Cameron, A. C., J. B. Gelbach, and D. L. Miller. 2008. "Bootstrap-based improvements for inference with clustered errors." *The Review of Economics and Statistics* 90 (3): 414–427.
- Carbonell-Nicolau, O., and E. A. Ok. 2007. "Voting over income taxation." *Journal of Economic Theory* 134 (1): 249–286.
- Carbonell-Nicolau, O., and E. F. Klor. 2003. "Representative democracy and marginal rate progressive income taxation." *Journal of Public Economics* 87 (9): 2339–2366.
- Cukierman, A., and A. H. Meltzer. 1988. "A political theory of progressive income taxation."
- Cusack, T., and P. Beramendi. 2006. "'Taxing work.'" *European Journal of Political Research* 45 (1): 43–73.

- De Donder, P., and J. Hindriks. 2003. "The politics of progressive income taxation with incentive effects." *Journal of Public Economics* 87 (11): 2491–2505.
- De Donder, P., and J. Hindriks. 2004. "Majority support for progressive income taxation with corner preferences." *Public Choice* 118 (3-4): 437–449.
- De Freitas, J. 2009. "A probabilistic voting model of progressive taxation with incentive effects." *Hacienda pública española* (190): 9–26.
- Dixit, A., and J. Londregan. 1996. "The determinants of success of special interests in redistributive politics." *Journal of Politics* 58: 1132–1155.
- Feldstein, M. S. 1969. "The effects of taxation on risk taking." *Journal of Political Economy*: 755–764.
- Feldstein, Martin, ed. 1983. *Behavioral Simulation Methods in Tax Policy Analysis*. University of Chicago Press.
- Figari, F., A. Paulus, and H. Sutherland. 2015. "Microsimulation and Policy Analysis." In *Handbook of Income Distribution*, ed. A. B. Atkinson and F. Bourguignon. Vol. 2. Elsevier.
- Förster, M., and M. Pearson. 2002. "Income distribution and poverty in the OECD area: trends and driving forces." *OECD Economic Studies* 34.
- Ganghof, S. 2006. "Tax mixes and the size of the welfare state: causal mechanisms and policy implications." *Journal of European Social Policy* 16 (4): 360–373.
- Holtz, E. D., W. Newey, and H. Rosen. 1988. "Estimating Vector Autoregression with Panel Data." *Econometrica* 56 (6): 1371–1395.
- International Labor Organization. 2008. *Income Inequalities in the Age of Financial Globalization*. Geneva: International Labor Office.
- Iversen, T., and D. Soskice. 2006. "Electoral institutions and the politics of coalitions: Why some democracies redistribute more than others." *American Political Science Review* 100 (2): 165–181.
- Kakwani, N. C. 1977a. "Applications of Lorenz curves in economic analysis." *Econometrica* 45 (3): 719–727.
- Kakwani, N. C. 1977b. "Measurement of tax progressivity: an international comparison." *The Economic Journal* 87 (345): 71–80.
- Kakwani, N. C., and N. Podder. 1976. "Efficient estimation of the Lorenz curve and associated inequality measures from grouped observations." *Econometrica* 44 (1): 137–148.
- Kato, J. 2003. *Regressive taxation and the welfare state: path dependence and policy diffusion*. Cambridge University Press.
- Kenworthy, L. 2008. "Taxes and Inequality: Lessons from Abroad." Last accessed June 12, 2018. <http://lanekenworthy.net/2008/02/10/taxes-and-inequality-lessons-from-abroad/>[URL].
- Kenworthy, L., and J. Pontusson. 2005. "Rising inequality and the politics of redistribution in affluent countries." *Perspectives on Politics* 3 (3): 449–471.
- Lambert, P. 2001. *The distribution and redistribution of income*. Manchester University Press.
- Lindbeck, A., and J. W. Weibull. 1987. "Balanced-budget redistribution as the outcome of political competition." *Public choice* 52 (3): 273–297.

- Lindbeck, A., and J. W. Weibull. 1993. "A model of political equilibrium in a representative democracy." *Journal of Public Economics* 51 (2): 195–209.
- Lindert, P. H. 2004. *Growing Public: Social Spending and Economic Growth since the Eighteenth Century, 2 volumes*. Cambridge University Press, New York.
- Marhuenda, F., and I. Ortuño-Ortín. 1995. "Popular support for progressive taxation." *Economics Letters* 48 (3): 319–324.
- Martin, C. J. 2015. "Labour market coordination and the evolution of tax regimes." *Socio-Economic Review* 13 (1): 33–54.
- Martin, I. W., A. K. Mehrotra, and M. Prasad. 2009. *The new fiscal sociology: Taxation in comparative and historical perspective*. Cambridge University Press.
- Newman, K. S., and R. O'Brian. 2011. *Taxing the Poor: Doing Damage to the Truly Disadvantaged*. Berkeley: University of California Press.
- Nickell, S. 1981. "Biases in dynamic models with fixed effects." *Econometrica* 49 (6): 1417–1426.
- OECD. 2007. *Benefits and Wages*. Paris: OECD.
- OECD. 2012. "Quality review of the OECD database on household incomes and poverty and the OECD earnings database." OECD, Social Policy Division.
- Persson, M. 1983. "The distribution of abilities and the progressive income tax." *Journal of Public Economics* 22 (1): 73–88.
- Pesaran, M. H. 2006. "Estimation and inference in large heterogeneous panels with a multifactor error structure." *Econometrica* 74 (4): 967–1012.
- Pesaran, M. H., and R. Smith. 1995. "Estimating long-run relationships from dynamic heterogeneous panels." *Journal of Econometrics* 68 (1): 79–113.
- Piketty, T., E. Saez, and S. Stantcheva. 2014. "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities." *American Economic Journal: Economic Policy* 6 (1): 230–71.
- Prasad, M., and Y. Deng. 2009. "Taxation and the worlds of welfare." *Socio-Economic Review* 7 (3): 431–457.
- Przeworski, A. 2010. *Democracy and the limits of self-government*. Vol. 9 Cambridge University Press.
- Rendon, S. R. 2012. "Fixed and Random Effects in Classical and Bayesian Regression." *Oxford Bulletin of Economics and Statistics* 75 (3): 460–476.
- Roemer, J. E. 1999. "The democratic political economy of progressive income taxation." *Econometrica*: 1–19.
- Roemer, J. E. 2011. "A theory of income taxation where politicians focus upon core and swing voters." *Social Choice and Welfare* 36 (3-4): 383–421.
- Rubin, D. 1996. "Multiple Imputation After 18+ Years." *Journal of the American Statistical Association* 91 (434): 473–489.
- Rubinstein, A. 1982. "Perfect equilibrium in a bargaining model." *Econometrica*: 97–109.
- Rueda, D., and D. Stegmueller. 2019. *Who Wants What?: Redistribution Preferences in Comparative Perspective*. Cambridge University Press.

- Shayo, M. 2009. "A Model of Social Identity with an Application to Political Economy: Nation, Class, and Redistribution." *American Political Science Review* 103 (2): 147–174.
- Shor, B., J. Bafumi, L. Keele, and D. Park. 2007. "A Bayesian multilevel modeling approach to time-series cross-sectional data." *Political Analysis* 15 (2): 165–181.
- Slitor, R. E. 1948. "The measurement of progressivity and built-in flexibility." *Quarterly Journal of Economics* 62 (2): 309–313.
- Snyder, J. M., and G. H. Kramer. 1988. "Fairness, self-interest, and the politics of the progressive income tax." *Journal of Public Economics* 36 (2): 197–230.
- Steinmo, S. 1989. "Political institutions and tax policy in the United States, Sweden, and Britain." *World Politics* 41 (04): 500–535.
- Steinmo, S. 1996. *Taxation and democracy: Swedish, British, and American approaches to financing the modern state*. New Haven: Yale University Press.
- Van Buuren, S. 2018. *Flexible imputation of missing data*. Boca Raton: CRC press.
- Zucman, G. 2015. *The hidden wealth of nations: The scourge of tax havens*. University of Chicago Press.

A. Proofs of Formal Statements

A.1. Proof of Lemma 1

Individuals choose t_M and t_R to maximize utility. Each individual in group R therefore solves:

$$\frac{\partial V_R}{\partial t_M} = p_M y_M + p_M t_M \frac{\partial y_M}{\partial t_M} = 0 \quad (19)$$

$$\frac{\partial V_R}{\partial t_R} = -y_R + p_R y_R + p_R t_R \frac{\partial y_R}{\partial t_R} = 0 \quad (20)$$

and each individual in groups M and P solve:

$$\frac{\partial V_M}{\partial t_M} = -y_M + p_M y_M + p_M t_M \frac{\partial y_M}{\partial t_M} = 0 \quad (21)$$

$$\frac{\partial V_M}{\partial t_R} = p_R y_R + p_R t_R \frac{\partial y_R}{\partial t_R} = 0 \quad (22)$$

and

$$\frac{\partial V_P}{\partial t_M} = p_M y_M + p_M t_M \frac{\partial y_M}{\partial t_M} = 0 \quad (23)$$

$$\frac{\partial V_P}{\partial t_R} = p_R y_R + p_R t_R \frac{\partial y_R}{\partial t_R} = 0. \quad (24)$$

From the first-order conditions in equations (19) and (20), type- R individuals choose $\hat{t}_M > 0$ to maximize tax revenue from type- M individuals, and because the condition $y_R > p_R y_R$ is always true, they prefer $\hat{t}_R = 0$.

From the first-order condition in equation (21), type- M individuals prefer $\hat{t}_M = 0$ because the condition $y_M > p_M y_M$ is always true. From the first-order condition in equation (22), type- M individuals also choose $\hat{t}_R > 0$ to maximize revenue from type- R individuals.

From the first-order conditions in equation (23) and (24), type- P individuals choose $\hat{t}_M, \hat{t}_R > 0$ to maximize revenue from type- M and - R individuals respectively.

Note that, using the elasticity expression in equation (6), we can express the revenue-maximizing tax rates on each group, expressed in equations (19) and (23) for group- M and equations (22) and (24) for group- R , as $\frac{t_M}{1-t_M} \eta_M = 1$ and $\frac{t_R}{1-t_R} \eta_R = 1$. Or, expressed solely in terms of η :

$$\frac{t_i}{1-t_i} = \frac{1}{\eta_i} \quad \text{or} \quad t_i = \frac{1}{1+\eta_i} \quad \text{for } i \in \{M, R\}$$

We can then conclude that preferences for tax progression are monotonically increasing and then decreasing in income. It follows directly from the preceding that $\hat{t}_R^M - 0 > \hat{t}_R^P - \hat{t}_M^P > 0 - \hat{t}_M^R$. Finally, that preferences for redistribution are decreasing in income follows from the fact that redistribution is increasing in t_M and t_R and that P has, relative to M , $\hat{t}_R^P = \hat{t}_R^M$ and $\hat{t}_M^P > \hat{t}_M^M = 0$, and M has, relative to R , $\hat{t}_R^M > \hat{t}_R^R = 0$ and $\hat{t}_M^R > \hat{t}_M^M = 0$. □

A.2. Proof of Proposition 1

From equation (10), the first-order conditions for party A's proposed platform are:

$$\frac{\partial \pi_A}{\partial t_R} = \sum_{i \in I} p_i f_i(\bar{\sigma}) g'(V_i) \frac{\partial V_i}{\partial t_R} = 0 \quad (25)$$

$$\frac{\partial \pi_A}{\partial t_M} = \sum_{i \in I} p_i f_i(\bar{\sigma}) g'(V_i) \frac{\partial V_i}{\partial t_M} = 0 \quad (26)$$

As can be easily checked, the first-order conditions for party B's platform are identical. This proves symmetry.

In the t_R case, we can write the first-order condition as:

$$-p_R f_R(\bar{\sigma}) g'(V_R) y_R + \sum p_i f_i(\bar{\sigma}) g'(V_i) \left[p_R y_R - p_R t_R \frac{\partial y_R}{\partial (1-t_R)} \right] = 0$$

Using the elasticity equation, this expression simplifies to:

$$\frac{p_R f_R(\bar{\sigma}) g'(V_R) y_R}{\sum p_i f_i(\bar{\sigma}) g'(V_i)} = p_R y_R \left[1 - \frac{t_R}{1-t_R} \eta_R \right]$$

Rearranging, we obtain an expression for the party's optimal choice of t_R :

$$\frac{t_R^*}{1-t_R^*} = \frac{1}{\eta_R} \left(1 - \frac{f_R(\bar{\sigma}) g'(V_R)}{\sum p_i f_i(\bar{\sigma}) g'(V_i)} \right) \quad (27)$$

Through identical steps, we can obtain a similar expression for the party's optimal choice of t_M :

$$\frac{t_M^*}{1-t_M^*} = \frac{1}{\eta_M} \left(1 - \frac{f_M(\bar{\sigma}) g'(V_M)}{\sum p_i f_i(\bar{\sigma}) g'(V_i)} \right) \quad (28)$$

Observe that a necessary and sufficient condition for the tax schedule to be progressive is:

$$\frac{t_R}{t_M} > 1 \quad \text{or} \quad \frac{t_R/(1-t_R)}{t_M/(1-t_M)} > 1$$

Dividing the right-hand side of equation (27) by the right-hand side of equation (28) this condition can be expressed as:

$$\frac{\eta_M}{\eta_R} \left(\frac{\sum p_i f_i(\bar{\sigma}) g'(V_i) - f_R(\bar{\sigma}) g'(V_R)}{\sum p_i f_i(\bar{\sigma}) g'(V_i) - f_M(\bar{\sigma}) g'(V_M)} \right) > 1 \quad (29)$$

It is easily observed that if $\eta_R = \eta_M$ and $f_R = f_M$, then the condition for the existence of a progressive tax schedule depends only on the existence of inequality, $V_R > V_M$. In particular, for $V_R > V_M$ and given the concavity of $g(\cdot)$, we have $g'(V_M) > g'(V_R)$, which makes the expression within parentheses in equation (29) larger than one.

□

A.3. Proof of Proposition 2

We show that with a mean-preserving spread in the income distribution, from X to Y , both redistribution and progressivity are lower under Y than X . Rewrite the first-order condition for a party's optimal choice of t_M , given by equation (26), as:

$$\sum_{i \in I \setminus M} p_i f_i(\bar{\sigma}) g'(V_i) \frac{\partial V_i}{\partial t_M} + p_M f_M(\bar{\sigma}) g'(V_M) \frac{\partial V_M}{\partial t_M} = 0 \quad (30)$$

We know from individual preferences of group- M members that they prefer $\hat{t}_M = 0$. Therefore, the utility of type- M individuals is decreasing for any positive level of t_M^* ; that is, for any $t_M^* > 0$, $\partial V_M / \partial t_M < 0$. Therefore, the second term in above equation (30) is negative and, in order for the first-order condition to be satisfied, the first term must be positive. This implies that for larger weights, p_R and p_P , on the portion of the first-order condition that is positive, and smaller weight, p_M , on the part of the first-order condition that is negative, the optimal choice of t_M moves closer to types- R and P ideal preferences. Therefore, for any $p_{X,P} < p_{Y,P}$, $p_{X,M} > p_{Y,M}$, and $p_{X,R} < p_{Y,R}$, we have $t_{M,Y}^* > t_{M,X}^*$.

In similar fashion, rewrite the first-order condition for a party's optimal choice of t_R , given from equation (25), as:

$$\sum_{i \in I \setminus M} p_i f_i(\bar{\sigma}) g'(V_i) \frac{\partial V_i}{\partial t_R} + p_M f_M(\bar{\sigma}) g'(V_R) \frac{\partial V_M}{\partial t_R} = 0 \quad (31)$$

From equation type- M preferences, given by equation (21), we know that type- M utility is increasing for any tax rate lower than that type's ideal (in any political equilibrium, such a rate will always prevail as long as type- R voters have any positive level of political influence, $p_R > 0$ and $f_R > 0$); that is, $\partial V_M / \partial t_R > 0$ for all $t_R \in [0, \hat{t}_R^M)$. Therefore, the second term in equation (31) being positive, the first term must be negative in order to satisfy the first-order condition. This implies that with larger weights on the negative part, and smaller weight on the positive portion, the equilibrium tax rate t_R falls. Therefore, for any $p_{X,P} < p_{Y,P}$, $p_{X,M} > p_{Y,M}$, and $p_{X,R} < p_{Y,R}$ we have $t_{R,Y}^* < t_{R,X}^*$.

Because the change shifts income upwards, and the drop in t_R does not compensate for the increase in t_M , (Lorenz curve) inequality in disposable income will increase, and therefore redistribution will decrease. \square

A.4. Proof of Proposition 3

Solving the model through backwards induction, we begin at the legislative bargaining stage. Following Binmore, Rubinstein, and Wolinsky (1986), we can use the Nash bargaining solution to model a more involved sequential strategic bargaining model (Rubinstein 1982) as the time period between offers goes to zero. Assume that party M is the formateur. If party M chooses party P to form a MP coalition, the solution to the bargaining problem is written as:

$$\mathbf{t}_{MP}^* = \arg \max\{[g(V_M(\mathbf{t}))][g(V_P(\mathbf{t}))]\} \quad (32)$$

The first-order conditions for this problem are:

$$\frac{1}{g(V_M(t_M, t_R))} g'(V_M) \frac{\partial V_M}{\partial t_M} + \frac{1}{g(V_P(t_M, t_R))} g'(V_P) \frac{\partial V_P}{\partial t_M} = 0 \quad (33)$$

$$\frac{1}{g(V_M(t_M, t_R))} g'(V_M) \frac{\partial V_M}{\partial t_R} + \frac{1}{g(V_P(t_M, t_R))} g'(V_P) \frac{\partial V_P}{\partial t_R} = 0 \quad (34)$$

These expressions are essentially averages of group-specific preferences, weighted by group-specific indirect utility. We already know that type- M and P individuals have identical preferences for t_R (from equations (22) and (24)). Therefore, $t_{MPR}^* = \hat{t}_R^P = \hat{t}_R^M$. Since $\hat{t}_M^P > \hat{t}_M^M = 0$, equation (33) implies that $t_{MP,M}^* \in (0, \hat{t}_M^P)$.

Now suppose that party M , as formateur, chooses party R to form a coalition. First-order conditions, similar to equations (33) and (34) obtain. Since, from group-specific preferences, neither M nor R prefers identical rates for either t_R or t_M : $\hat{t}_R^M > \hat{t}_R^R = 0$ and $\hat{t}_M^R > \hat{t}_M^M = 0$. It follows that tax rates bargained by the coalition will be such that $t_{MR,R}^* \in (0, \hat{t}_R^M)$ and $t_{MR,M}^* \in (0, \hat{t}_M^R)$. Note also that the bargained tax rate on type- M individuals is lower in a MR coalition than a MP coalition: $t_{MR,M}^* < t_{MP,M}^*$. The first-order condition for t_M in a MR coalition is:

$$\frac{1}{g(V_M(t_M, t_R))} g'(V_M) \frac{\partial V_M}{\partial t_M} + \frac{1}{g(V_R(t_M, t_R))} g'(V_R) \frac{\partial V_R}{\partial t_M} = 0 \quad (35)$$

Compare this to the first-order condition for the t_M tax rate in an MP coalition, given by equation (33). From type- R and type- P individual preferences for t_M , given respectively by equations (19) and (23), we can see that $\partial V_R / \partial t_M = \partial V_P / \partial t_M$. Hence, evaluating (35) at $t_{MR,M}^*$ makes that expression negative since $1/g(V_M) > 1/g(V_P)$, $g'(V_P) > g'(V_R)$, and $\partial V_M / \partial t_M < 0$ for all $t_M \in [0, 1]$. Thus, in order to satisfy the first-order condition (35), t_M needs to become smaller, relative to $t_{MR,M}^*$.

By similar reasoning, we can establish equilibrium values for a PR coalition. Since types- P and R have identical preferences for t_M (from, again, equations (19) and (23)) we have $t_{PR,M}^* = \hat{t}_M^P = \hat{t}_M^R$. For t_R type- R individuals of course prefer $\hat{t}_R^R = 0$ so $t_{PR,R}^* \in [0, \hat{t}_R^P]$. Note that, because parties R and P both prefer a revenue maximizing tax rate t_M but that the bargained tax rate $t_{PR,R}^*$ is less than the revenue maximizing tax rate t_R , the tax schedule resulting from a PR coalition could possibly be regressive.

Given the equilibrium values of tax rates for each kind of coalition, we can solve for which coalitions will form, for a given formateur. The critical result is that, for $p_R y_R$ large enough, M prefers to form a coalition with P , P prefers to form a coalition with M , and R prefers to form a coalition with P . The reasoning is as follows. Parties M and P want a revenue-maximizing tax rate on type- R citizens, but in a coalition with P , M must accept a positive tax rate on its own group. If M forms a coalition with R it can get a lower rate on its own group, but still higher than it prefers, but must also accept a less-than-maximizing rate on the rich. Put simply, if the rich are rich enough, the revenue lost from a lower rate on the rich will swamp the income retained by lowering the tax rate slightly on its own group. Similarly, for party P , parties P and M both want to soak the rich while party P in a MP coalition has to settle for a lower rate on the middle class than it prefers. At the same time, parties P and R both want to soak the middle class while party P

in a PR coalition has to settle for a lower rate on the rich than it prefers. When the rich are richer, party P prefers to soak the rich rather than the middle class. Finally, since both party R and party P want a revenue-maximizing rate on the middle class, but there is no consensus on either rate between parties R and M , party R prefers a coalition with party P .

Given the equilibrium tax rates and the expected coalitions that will form, we can analyze voters' decisions. Given the assumptions made in the text, the results follow quite simply. First, note that if votes were allocated solely by the symmetrically-distributed ideological preferences, each party would receive the same share of votes: exactly one-third of the votes from each group in the population. But this will not be the case because of preference proximity for policy between type- P and $-M$ voters. To see this, observe that whether party- P or party- M is formateur, the policy outcome will be identical, since P prefers to form a coalition with M and vice versa. Hence, the indifferent voter in groups M and P is ideological neutral. For example, for a type- P voter:

$$g(V_P(\mathbf{t}_M)) - g(V_P(\mathbf{t}_P)) = \bar{\sigma}_{MP} = 0.$$

By contrast, the indifferent voter choosing between party- R and either party- M or $-P$ must have a strictly positive ideological preference for party- R since policy $\mathbf{t}_M = \mathbf{t}_P$ gives strictly higher economic utility than policy \mathbf{t}_R . Again, for a type- P voter, for example:

$$g(V_P(\mathbf{t}_M)) - g(V_P(\mathbf{t}_R)) = \bar{\sigma}_{PR} > 0.$$

Therefore, more than a third of type- R voters will vote for party R while less than a third of type- P and $-M$ voters will vote for party R . Conversely, more than a third of type- P and $-M$ voters but less than a third of type- R voters will vote for parties P and M . Hence, because group- R constitutes strictly less than half of the population, parties P and M will get a larger share of votes than party R , making one of those two parties the formateur with certainty.

Hence, either party P or M will be the formateur, and the legislative coalition will adopt the policy \mathbf{t}_{MP} . Compared to the policy chosen in a majoritarian system, $\hat{\mathbf{t}}_M$, the \mathbf{t}_{MP} policy is less progressive. The tax rate on the rich, t_R , is identical under each policy, but t_M is higher under \mathbf{t}_{MP} , reflecting the preferences of the poor, than it is under a majoritarian system, which caters to the preferences of the middle class. The proportional representation policy, \mathbf{t}_{MP} , is also more redistributive compared to the majoritarian policy, $\hat{\mathbf{t}}_M$, since it increases the disposable income of the poor and, consequently, reduces income inequality.

□

B. Empirical details

Table B.1 shows countries and years included in our analysis. In many cases, the limiting factor is information on inequality indices needed to calculate our measure of redistribution. Note that we conduct a robustness using multiple imputation (assuming that the process leading to missing inequality information in a given year is MAR) and found no substantive difference in results (see Table II).

Table B.1.
Countries and years included in our analysis

Country	Years included in analysis
Australia	2004, 2008, 2010, 2012, 2014
Austria	2007–2015
Belgium	2004–2015
Canada	2001–2015
Denmark	2005–2014
Finland	2001–2015
France	2005, 2008, 2009–2015
Germany	2004, 2008, 2009–2014
Greece	2004–2015
Iceland	2004–2014
Ireland	2005–2014
Japan	2003, 2006, 2009, 2012
Netherlands	2005–2014
New Zealand	2003, 2008, 2009, 2011, 2012, 2014
Norway	2004, 2008, 2009–2014
Portugal	2004–2015
Spain	2007–2015
Sweden	2004, 2008, 2009–2015
Switzerland	2009, 2011, 2013, 2014
United Kingdom	2001–2015
United States	2005, 2008, 2009–2015

Table B.2 shows average tax function parameter estimates (across all years). Besides the clear differences in tax structures between countries, it also shows substantial over-time variation *within* countries: while the pooled standard deviation for τ is 4.8, the within country standard deviation is 1.6; for λ the overall standard deviation is 4.6 with an within-country SD of 2. We employ this within-country variation in our empirical analysis.

Figure B.1 shows the distribution of tax and benefit progressivity in our sample of 21 OECD countries (for 2001 to 2015).

Figure B.2 plots τ against the flat-rate tax parameter. Superimposed is a LOESS smoother with 95% confidence bands (red line) and a quadratic least-squares fit with 95% confidence bands (green line).

Table B.2.
Summary of estimated tax function parameters

Country	τ [$\times 100$]	λ
Australia	17.73	5.51
Austria	17.38	5.53
Belgium	21.96	8.79
Canada	19.60	7.52
Denmark	21.23	10.78
Finland	14.82	3.61
France	6.71	1.85
Germany	15.12	4.53
Greece	19.83	7.15
Iceland	19.88	19.97
Ireland	17.99	6.00
Japan	7.98	3.22
Netherlands	24.45	11.49
New Zealand	10.51	2.48
Norway	16.40	6.72
Portugal	12.49	3.25
Spain	13.96	3.84
Sweden	19.76	9.07
Switzerland	13.03	4.85
United Kingdom	13.76	3.49
United States	10.88	2.89
Pooled mean	15.97	6.31
Pooled std.dev.	4.84	4.58
Within-country std.dev.	1.55	2.14

Note: Parameter estimates of equation 12, 2001–2015 averages.
 Within-country std.dev. calculated on $\tau_{it} - \bar{\tau}_i + \bar{\tau}$ (*mutatis mutandis* for λ).

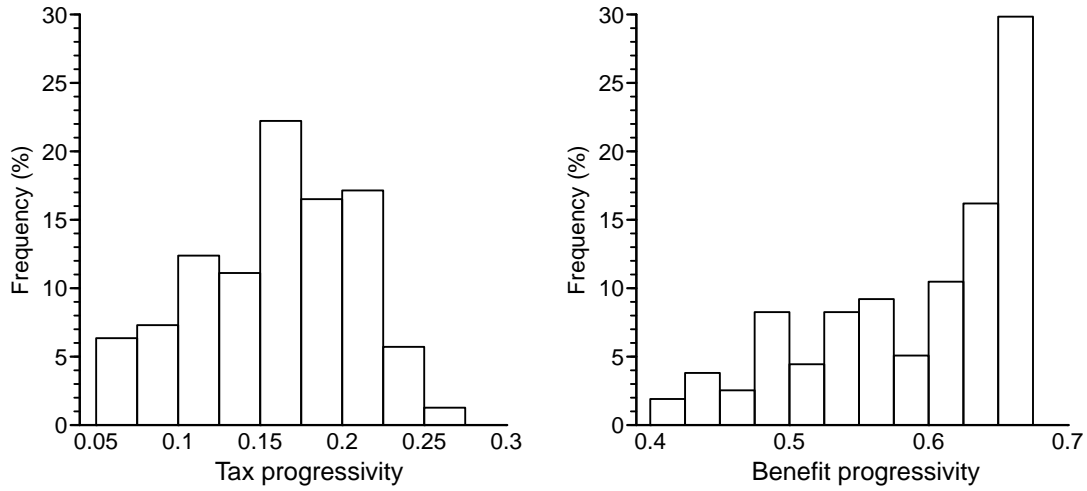


Figure B.1.
Distribution of tax and benefit progressivity in 21 OECD countries, 2001-2015

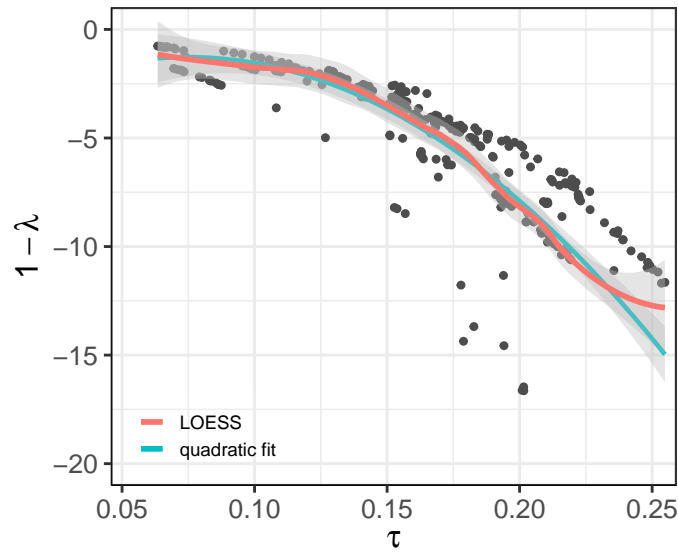


Figure B.2.
The relationship between the flat-rate tax parameter and progressivity

This figure plots tax function parameter estimates for the flat-rate tax parameter $1 - \lambda$ against progressivity parameter estimates, τ , for 21 OECD countries from 2001 to 2015. Superimposed are a LOESS smoother (with span 1/2) and a quadratic linear model fit with 95% confidence bands.

C. Low income turnout in 21 OECD countries

As measure of unequal influence of lower income individuals we calculate the elasticity of an individuals turnout to his or her income. Since we need to cover 21 OECD countries, we use data from modules 2 to 4 (2001-2016) of the Comparative Study of Electoral Systems (CSES).

While the CSES provides broad coverage, this does come at the cost of some harmonization issues. By design income in each country should be measured by placing respondents in income quintiles of their annual pre-tax post-transfer household income. There are some country-specific deviations from this common scheme. We exclude surveys where the income concept used was explicitly post-tax (for example, the Netherlands in its 2006 module uses “disposable income”). Where surveys asked for monthly instead of annual amount, the CSES transformed them to annual income figures. A more difficult problem to address is the fact that some countries do not report income quintiles but groupings that likely are based on country-specific domain knowledge. To define *low income* respondents we use the bottom two country quintiles or the bottom two country-specific groups. Turnout in the CSES is captured by a question probing if respondents voted in the last lower house election or (the first round of) the presidential election. According to general CSES guidelines question wording should minimize over-reporting, but the actual wording follows national election survey standards. In some countries non-voters are simply those that selected a ‘did not vote’ category in a list of party choices. Under compulsory voting rules (in the Australian case) voters who reported to have voted “informally” or did not vote where counted as not having voted. We exclude all respondents who are (self-reported) ineligible to vote or below voting age. The resulting data set provides us with at least one surveys per country with a minimum sample size of 1,574 cases. The median sample size is 5,899.

Any analysis of the relationship between turnout and income has to confront the problem of item-nonresponse. The average percentage of missing values for turnout in our sample is 3.3, but for some countries turnout non-response is higher (10.8% in the US and 6.7% in Canada). Respondents are less likely to state their household income: about 16% of our CSES sample has missing household income information ranging from 3% in the Netherlands to 37% in Spain. Respondents with missing income are also more likely to have missing values for turnout (the mean difference in share of missing values is 0.016 with a two sided-p value of 0.000). While there is no straightforward solution for this issue we argue that assuming that responses are missing completely at random (Rubin 1996)—as assumed when using listwise deletion—is the least appropriate approach. We assume that responses are missing at random conditional on a set of predictive covariates for turnout and income. For the latter we specify a Mincer-type equation imputing income as a function of education-specific age-income curves. We impute missing values using chained regression equations (Van Buuren 2018). After iterating the imputation chains for 20 iterations, we generated 10 imputed data sets. All calculations reported below are based on these 10 imputations. The imputation equation for turnout is a logit model with a second order polynomial of age, an indicator for tertiary education (BA or equivalent and beyond), an indicator variable for gender and an indicator equal to one if at least one person in the household is a union member. Income is treated as a set of ordered categories imputed using an ordered logit specification. We use a Mincer equation of quadratic age-education curves for the respondents augmented with indicators for the respondent’s gender and the union membership status of the household.

With the imputed data sets in hand we calculate for each country the turnout rate of individuals in the bottom two income quintiles (or groups). We weight for sample inclusion probabilities where available.